

SEQUENCE LISTING

<110> JACK, WILLIAM E.
GARDNER, ANDREW
NEW ENGLAND BIOLABS, INC.

<120> INCORPORATION OF MODIFIED NUCLEOTIDES BY ARCHAEON DNA
POLYMERASES AND RELATED METHODS

<130> NEB-166-PCT

<140>

<141>

<150> 60/157,204

<151> 1999-09-30

<160> 30

<170> PatentIn Ver. 2.0

<210> 1

<211> 36

<212> DNA

<213> Synthetic Oligonucleotides

<400> 1

caggcagagg cttataaaaa tcctcgccaa cagctt

36

<210> 2

<211> 26

<212> DNA

<213> Synthetic Oligonucleotides

<400> 2

ggtggcagca gccaaactcag cttcct

26

<210> 3

<211> 24

<212> DNA

<213> Synthetic Oligonucleotides

<400> 3

gattctcatg ataagctacg ccga

24

<210> 4

<211> 5837

<212> DNA

<213> *Thermococcus litoralis*

<400> 4

```

gaattcgcgga taaaatctat tttcttcctc catttttcaa tttcaaaaac gtaagcatga 60
gccaaacctc tcgccctttc tctgtccttc ccgctaacc ccttgaaaac tctctccaaa 120
gcattttttg atgaaagctc acgctcctct atgagggtca gtatatctgc aatgagttcg 180
tgaagggtta ttctgtagaa caactccatg attttcgatt tggatggggg tttaaaaatt 240
tggcggaact tttatttaat ttgaactcca gtttatactt ggtggtattt atgatactgg 300
acactgatta cataacaaaa gatggcaagc ctataatccg aatttttaag aaagagaacg 360
gggagtttaa aatagaactt gaccctcatt ttcagcccta tatatatgct cttctcaaag 420
atgactccgc tattgaggag ataaaggcaa taaagggcga gagacatgga aaaactgtga 480
gagtgtctga tgcagtgaat gtcaggaaaa aatttttggg aagggaagtt gaagtctgga 540
agctcatttt cgagcatccc caagacgttc cagctatgcg gggcaaaata agggaacatc 600
cagctgtggt tgacatttac gaatatgaca taccctttgc caagcgttat ctcatagaca 660
agggcttgat tcccatggag ggagacgagg agcttaagct ccttgccttt gatattgaaa 720
cgttttatca tgaggagat gaatttgga agggcgagat aataatgatt agttatgccg 780
atgaagaaga ggccagagta atcacatgga aaaatatcga tttgccgtat gtcgatgttg 840
tgtccaatga aagagaaatg ataaagcgtt ttgttcaagt tgttaaagaa aaagaccccg 900
atgtgataat aacttacaat ggggacaatt ttgatttgcc gtatctcata aaacgggcag 960
aaaagctggg agttcggcct gtcttaggaa gggacaaaaga acatcccgaa cccaagattc 1020
agaggatggg tgatagtttt gctgtggaaa tcaagggtag aatccacttt gatcttttcc 1080
cagttgtgcg aaggacgata aacctcccaa cgtatacgct tgaggcagtt tatgaagcag 1140
ttttaggaaa aaccaaaagc aaattaggag cagaggaaat tgccgctata tgggaaacag 1200
aagaaagcat gaaaaaacta gccagtagt caatggaaga tgctagggca acgtatgagc 1260
tcgggaagga attcttcccc atggaagctg agctggcaaa gctgataggt caaagtgtat 1320
gggacgtctc gagatcaagc accggcaacc tcgtggagtg gtatctttta aggggtggcat 1380
acgcgaggaa tgaacttgca ccgaacaaac ctgatgagga agagtataaa cggcgcttaa 1440
gaacaactta cctgggagga tatgtaaaag agccagaaaa aggtttgtgg gaaaatatca 1500
tttatttgga tttccgcagt ctgtaccctt caataatagt tactcacaac gtatccccag 1560
atacccttga aaaagagggc tgtaagaatt acgatgttgc tccgatagta ggatataggt 1620
tctgcaagga ctttcggggc tttattccct ccatactcgg ggacttaatt gcaatgaggc 1680
aagatataaa gaagaaaatg aaatccacaa ttgaccgat cgaaaagaaa atgctcgatt 1740
ataggcaaaag ggctattaaa ttgcttgcaa acagcatctt acccaacgag tggttaccaa 1800
taattgaaaa tggagaaata aaattcgtga aaattggcga gtttataaac tcttacatgg 1860
aaaaacagaa ggaaaacgtt aaaacagtag agaatactga agttctcgaa gtaaacaacc 1920
tttttgcat ctcattcaac aaaaaaatca aagaaagtga agtcaaaaaa gtcaaagccc 1980
tcataagaca taagtataaa gggaaagcct atgagattca gcttagctct ggtagaaaaa 2040
ttaacataac tgctggccat agtctgttta cagttagaaa tggagaaata aaggagttt 2100
ctggagatgg gataaaagaa ggtgacctta ttgtagcacc aaagaaaatt aaactcaatg 2160
aaaaaggggt aagcataaac attcccgagt taatctcaga tctttccgag gaagaaacag 2220
ccgacattgt gatgacgatt tcagccaagg gcagaaagaa cttctttaaa ggaatgctga 2280
gaactttaag gtggatgttt ggagaagaaa atagaaggat aagaacattt aatcgctatt 2340
tgttccatct cgaaaaacta ggccttatca aactactgcc ccgcggatat gaagttactg 2400
actgggagag attaaagaaa tataaacaac tttacgagaa gcttgctgga agcgttaagt 2460
acaacggaaa caagagagag tatttagtaa tgttcaacga gatcaaggat tttatatctt 2520
acttcccaca aaaagagctc gaagaatgga aaattggaac tctcaatggc tttagaacga 2580
attgtattct caaagtcgat gaggattttg ggaagctcct aggttactat gttagtgagg 2640
gctatgcagg tgcacaaaaa aataaaactg gtggtatcag ttattcggtg aagctttaca 2700

```

```

atgaggaccc taatgttctt gagagcatga aaaatgttgc agaaaaattc tttggcaagg 2760
ttagagttga cagaaattgc gtaagtatat caaagaagat ggcatactta gttatgaaat 2820
gcctctgtgg agcattagcc gaaaacaaga gaattccttc tgttatactc acctctcccg 2880
aaccggtacg gtggtcattt ttagaggcgt attttacagg cgatggagat atacatccat 2940
caaaaagggt taggctctca acaaaaagcg agctccttgc aaatcagctt gtgttcttgc 3000
tgaactcttt gggaatatcc tctgtaaaga taggctttga cagtggggtc tatagagtgt 3060
atataaatga agacctgcaa tttccacaaa cgtctaggga gaaaaacaca tactactcta 3120
acttaattcc caaagagatc cttagggacg tgtttggaag agagttccaa aagaacatga 3180
cgttcaagaa atttaaagag cttgttgact ctggaaaact taacagggag aaagccaagc 3240
tcttgagatt cttcattaat ggagatattg tccttgacag agtcaaaagt gttaaagaaa 3300
aggactatga agggatgtc tatgacctaa gcgttgagga taacgagAAC tttcttggtg 3360
gttttggttt gctctatgct cacaacagct attacggcta tatggggtat cctaaggcaa 3420
gatggtactc gaaggaatgt gctgaaagcg ttaccgcatg ggggagacac tacatagaga 3480
tgacgataag agaaatagag gaaaagttcg gctttaaggt tctttatgcg gacagtgtct 3540
caggagaaag tgagatcata ataaggcaaa acggaaagat tagattttgtg aaaataaagg 3600
atcttttctc taagggtggac tacagcattg gcgaaaaaga atactgcatt ctcgaagggtg 3660
ttgaagcact aactctggac gatgacggaa agcttgtctg gaagcccgtc ccctacgtga 3720
tgaggcacag agcgaataaa agaattgtcc gcactctggc gaccaacagc tggatatag 3780
atgttactga ggatcattct ctcataggct atctaaacac gtcaaaaacg aaaactgcca 3840
aaaaaatcgg ggaaagacta aaggaagtaa agccttttga attaggcaaa gcagtaaaat 3900
cgctcatatg cccaaatgca ccgttaaagg atgagaatac caaaactagc gaaatagcag 3960
taaaattctg ggagctcgta ggattgattg taggagatgg aaactggggg ggagattctc 4020
gttgggcaga gtattatctt ggactttcaa caggcaaaga tgcagaagag ataaagcaaa 4080
aacttctgga acccctaaaa acttatggag taatctcaaa ctattacca aaaaacgaga 4140
aaggggactt caacatcttg gcaaagagcc ttgtaaagtt tatgaaaagg cactttaagg 4200
acgaaaaagg aagacgaaaa attccagagt tcatgtatga gcttccggtt acttacatag 4260
aggcatttct acgaggactg ttttcagctg atgggtactgt aactatcagg aaggggattc 4320
cagagatcag gctaacaaac attgatgctg acttttctaag ggaagtaagg aagcttctgt 4380
ggattgttgg aatttcaaatt tcaatatttg ctgagactac tccaaatcgc tacaatggtg 4440
tttctactgg aacctactca aagcatctaa ggatcaaaaa taagtggcgt tttgctgaaa 4500
ggataggctt tttaatcgag agaaagcaga agagactttt agaacattta aaatcagcga 4560
gggtaaaaag gaataccata gattttggct ttgatcttgt gcatgtgaaa aaagtcgaag 4620
agataccata cgagggttac gtttatgaca ttgaagtcca agagacgcat aggttctttg 4680
caaacaacat cctggtacac aatactgacg gcttttatgc cacaataccc ggggaaaagc 4740
ctgaactcat taaaaagaaa gccaaggaat tctaaacta cataaactcc aaacttccag 4800
gtctgcttga gcttgagtat gagggctttt acttgagagg attctttgtt acaaaaaagc 4860
gctatgcagt catagatgaa gagggcagga taacaacaag gggcttggaa gtagtaagga 4920
gagattggag tgagatagct aaggagactc aggcaaagggt ttagaggctt atacttaaag 4980
agggaaagtgt tgaaaaagct gtagaagttg ttagagatgt ttagagaaa atagcaaaat 5040
acagggttcc acttgaaaag cttgttatcc atgagcagat taccagggat ttaaaggact 5100
acaaagccat tggccctcat gtcgcgatag caaaaagact tgccgcaaga gggataaaaag 5160
tgaaaccggg cacaataata agctatatcg ttctcaaagg gagcggaag ataagcgata 5220
gggtaatttt acttacagaa tacgatccta gaaaacacaa gtacgatccg gactactaca 5280
tagaaaacca agttttgccg gcagtactta ggatactcga agcgtttgga tacagaaagg 5340
aggatttaag gtatcaaagc tcaaaacaaa ccggcttaga tgcattggctc aagaggtagc 5400
tctgttgctt tttagtccaa gtttctccgc gagtctctct atctctcttt tgtattctgc 5460
tatgtggttt tcattcacta ttaagtagtc cgccaaagcc ataacgcttc caattccaaa 5520
cttgagctct ttccagtctc tggcctcaaa ttcactccat gtttttggat cgtcgttct 5580

```

ccctcttctg ctaagcctct cgaatctttt tcttggcgaa gagtgtacag ctatgatgat 5640
tatctcttcc tctggaaacg catcttttaa cgtctgaatt tcacttagag acctcactcc 5700
gtcgattata actgccttgt acttcttttag tagttctttt acctttggga tcgttaattt 5760
tgccacggca ttgtcccca gtcctgcct aagctgaatg ctcacactgt tcataccttc 5820
gggagttctt gggatcc 5837

<210> 5

<211> 15

<212> PRT

<213> *Thermococcus litoralis*

<400> 5

Ala Ile Lys Leu Leu Ala Asn Ser Tyr Tyr Gly Tyr Met Gly Tyr
1 5 10 15

<210> 6

<211> 15

<212> PRT

<213> *Pyrococcus* Sp. (GB-D)

<400> 6

Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 7

<211> 15

<212> PRT

<213> *Thermococcus* sp.

<400> 7

Ala Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 8

<211> 15

<212> PRT

<213> *Pyrococcus furiosus*

<400> 8

Ala Ile Lys Leu Leu Ala Asn Ser Phe Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 9

<211> 15

<212> PRT

<213> Thermococcus fumicolans

<400> 9

Ala Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 10

<211> 15

<212> PRT

<213> Thermococcus gorgonarius

<400> 10

Ala Ile Lys Ile Leu Ala Asn Ser Phe Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 11

<211> 15

<212> PRT

<213> Thermococcus sp. (TY)

<400> 11

Ala Val Lys Leu Leu Ala Asn Ser Tyr Tyr Gly Tyr Met Gly Tyr
1 5 10 15

<210> 12

<211> 15

<212> PRT

<213> Pyrococcus abyssi

<400> 12

Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 13

<211> 15

<212> PRT

<213> Pyrococcus glycovaorans

<400> 13

Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 14
<211> 15
<212> PRT
<213> *Pyrococcus horikoshii*

<400> 14
Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 15
<211> 15
<212> PRT
<213> *Pyrococcus* sp. (GE23)

<400> 15
Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 16
<211> 15
<212> PRT
<213> *Pyrococcus* Sp. (KOD1)

<400> 16
Ala Ile Lys Ile Leu Ala Asn Ser Tyr Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 17
<211> 15
<212> PRT
<213> *Pyrococcus woesei*

<400> 17
Ala Ile Lys Leu Leu Ala Asn Ser Phe Tyr Gly Tyr Tyr Gly Tyr
1 5 10 15

<210> 18
<211> 15
<212> PRT
<213> *Archaeoglobus fulgidus*

<400> 18
Thr Leu Lys Val Leu Thr Asn Ser Phe Tyr Gly Tyr Met Gly Trp
1 5 10 15

<210> 19

<211> 15

<212> PRT

<213> *Cenarchaeum symbiosum*

<400> 19

Ala Leu Lys Val Val Leu Asn Ala Ser Tyr Gly Val Met Gly Ala
1 5 10 15

<210> 20

<211> 15

<212> PRT

<213> *Methanococcus jannaschii*

<400> 20

Ser Ile Lys Ile Leu Ala Asn Ser Val Tyr Gly Tyr Leu Ala Phe
1 5 10 15

<210> 21

<211> 15

<212> PRT

<213> *Methanococcus voltae*

<400> 21

Ser Ile Lys Val Leu Ala Asn Ser His Tyr Gly Tyr Leu Ala Phe
1 5 10 15

<210> 22

<211> 15

<212> PRT

<213> *Pyrodictium occultum*

<400> 22

Ala Leu Lys Val Leu Ala Asn Ala Ser Tyr Gly Tyr Met Gly Trp
1 5 10 15

<210> 23

<211> 15

<212> PRT

<213> *Sulfurisphaera ohwakuensis*

<400> 23

Ala Met Lys Val Phe Ile Asn Ala Thr Tyr Gly Val Phe Gly Ala
1 5 10 15

<210> 24

<211> 15

<212> PRT

<213> Sulfolobus acidocaldarius

<400> 24

Ala Met Lys Val Phe Ile Asn Ala Thr Tyr Gly Val Phe Gly Ala
1 5 10 15

<210> 25

<211> 15

<212> PRT

<213> Sulfolobus solfataricus

<400> 25

Ala Met Lys Val Phe Ile Asn Ala Thr Tyr Gly Val Phe Gly Ala
1 5 10 15

<210> 26

<211> 15

<212> PRT

<213> Herpesvirus

<400> 26

Ala Ile Lys Val Val Cys Asn Ser Val Tyr Gly Phe Thr Gly Val
1 5 10 15

<210> 27

<211> 15

<212> PRT

<213> human herpesvirus 2

<400> 27

Ala Ile Lys Val Val Cys Asn Ser Val Tyr Gly Phe Thr Gly Val
1 5 10 15

<210> 28

<211> 15

<212> PRT

<213> Human cytomegalovirus

<400> 28

Ala Leu Lys Val Thr Cys Asn Ala Phe Tyr Gly Phe Thr Gly Val
1 5 10 15

<210> 29

<211> 15

<212> PRT

<213> Human DNA Polymerase alpha

<400> 29

Ala Leu Lys Leu Thr Ala Asn Ser Met Tyr Gly Cys Leu Gly Phe
1 5 10 15

<210> 30

<211> 15

<212> PRT

<213> Phage T4

<400> 30

Asn Arg Lys Ile Leu Ile Asn Ser Leu Tyr Gly Ala Leu Gly Asn
1 5 10 15